

A NOTE ON MOLECULAR CONFIGURATION IN SODIUM THYMONUCLEATE

Rosalind E. Franklin and R. G. Gosling

17/3/53.

Sodium thymonucleate fibres give two distinct types of X-ray diagram. The first, corresponding to a crystalline form obtained at about 75% relative humidity, has been described in detail elsewhere (). At high^{er} humidities a new structure, showing a lower degree of order appears, and persists over a wide range of ambient humidity and water content. The water content of the fibres, which are crystalline at lower humidities, may vary from about 50% to several hundred per cent. of the dry weight in this structure. Other fibres which do not give crystalline structure at all, show this less ordered structure at much lower humidities. The diagram of this structure, which we have called structure B, shows in striking manner the features characteristic of helical structures (). Although this cannot be taken as proof that the structure is helical, other considerations make the existence of a helical structure highly probable.

Structure B is derived from the crystalline structure A when the sodium thymonucleate (NaDNA) fibres take up quantities of water in excess of about 40% of their weight. The change is accompanied by an increase of about 30% in the length of the fibre, and by a substantial re-arrangement of the molecule. It therefore seems reasonable to assume that in structure B the structural units of NaDNA (molecules or groups of molecules) are relatively free from the influence of neighbouring groups, each unit being shielded by a sheath of water. Each unit is therefore free to take up its least-energy configuration independently of its neighbours and, in view of the nature of the long-chain molecules involved, it is highly likely that the

general form will be helical (). If we adopt the hypothesis of a helical structure, it is immediately possible, from the X-ray diagram of structure B, to make certain deductions as to the nature and dimensions of the helix.

From the angle between the straight lines which can be drawn through the origin and the innermost maxima of the 1st, 2nd, 3rd and 5th layer-lines, the diameter of the helix can be calculated. It is found to be about 20A. Since this linear array of maxima is one of the strongest features of the diagram, we must conclude that a (crystallographically) very important part of the molecule lies on a helix of this diameter. This can only be the phosphate groups (or, perhaps, the phosphorus atoms). Thus, if the structure is helical we find that the phosphate groups lie on a helix of diameter about 20A, and the sugar and base groups must accordingly be turned inwards towards the helical axis.

This is in agreement with the conclusion which we reached previously by quite other reasoning (), namely that, whatever the structural unit, the phosphate groups must be on the outside. There were two principal reasons for believing this. The first derives from the work of Gulland and his collaborators who showed that even in aqueous solution the -CO and -NH₂ groups of the bases are inaccessible and cannot be titrated, whereas the phosphate groups are fully accessible. The second is our own observations on the way in which the structural units in the crystalline structure A are floated apart by an excess of water, the process being a continuous one which leads to the formation first of a gel and ultimately to a solution. The hygroscopic part of the molecule may be presumed to lie in the phosphate groups; ((C₂H₅O)₂PO₂Na and (C₃H₇O)₂PO₂Na are highly hygroscopic) and the simplest explanation of the above process is that these groups lie on the outside of the structural units. Furthermore the ready availability of the phosphate groups for interaction with proteins can also be explained this way.

The above estimate of 20A diameter was based on the assumption of a single strand helix. That is, the first maximum on the n^{th} layer-line corresponds to the first maximum in $J_n(2\pi rR)$. Where $J_n(u)$ is the n^{th} order Bessel Function of u , r is the radius of the helix and R the distance from the fibre-axis direction in reciprocal space.

The strong meridional maximum at 3.4 \AA° lies accurately on the 10th layer-line. From this new lines of maxima emanate, as from the origin, crossing the origin series on the 5th layer-line, corresponding to a $J_5(u)$ for each series, confirming that the second origin does lie on the 10th layer-line. This then, indicates that there are 10 structural units in one turn of the single-strand helix. For a helix of diameter 20 \AA° this gives a distance of 6A between neighbouring units in one molecule, which is a reasonable distance for the P-P value in NaDNA. (this distance in a fully extended chain is 6.8 \AA°).

If, instead of a single-strand helix we propose 2 equally spaced co-axial helical molecules, the first maximum on the n^{th} layer-line corresponds to the first maximum in $J_{2n}(2\pi rR)$. Since our value of R is fixed and the first maximum in $J_2(x)$ occurs at very nearly twice the value in x of the first maximum in $J_1(x)$ (which gave us $2r \approx 20 \text{ \AA}^\circ$) the value of $2r$ for a 2-strand helix must be $\approx 40 \text{ \AA}^\circ$. The cross-section of the helix would then be considerably greater than that of the primitive cell in the crystalline structure A, and this would seem highly improbable. The same argument, with even more force, eliminates the possibility of 3 equally spaced co-axial helical molecules.

On the theory of a single-strand helix, the series of equatorial maxima should correspond to the maxima of $J_0(4\pi r \sin \theta)$. The maxima on our photograph do not, however, fit this function. This is rather to be expected. For we know that the helix so far considered is only the most important member of a series of co-axial helices of different radii, the non-phosphates parts

of the molecule must lie on a series of co-axial helicies of smaller radii. Following Crick, Cochran and Vand, the structure factor on the n th layer-line for a series of co-axial helicies

$$F_n = \sum_j f_j J_n(2\pi R r_j) e^{i[n(\psi - \phi_j + \frac{1}{2}\pi) + 2\pi \frac{y_j}{c}]}]$$

(here give definitions)

Simplifying this, for the case of a whole number of residues per turn of the helix, we readily obtain

$$F_n = e^{in(\psi + \pi/2)} [f_1 J_n(x_1) e^{in\alpha_1} + f_2 J_n(x_2) e^{in\alpha_2} + \dots + f_j J_n(x_j) e^{in\alpha_j}]$$

where $\alpha_k = \phi_k - 2\pi y_k/c$

and $x_k = 2\pi R r_k$.

It follows that

$$I = \sum_j f_j^2 [J_n(x_j)]^2 + \sum_{j/k}^{jk} f_j f_k J_n(x_j) J_n(x_k) \cos [n(\alpha_j - \alpha_k)]$$

From this it is evident that the innermost maxima on the layer-lines will always be given by the helix of the largest diameter, containing the terms ~~existing~~ $J_n(x_j) J_n(x_k)$ for values of x_j and x_k smaller than the maximum being, in the region of the first maxima, very small. Later maxima, however, may be obliterated or shifted owing to the appearance of important negative terms in the expression for I .

Thus, while we do not attempt to offer a complete interpre-

tation of the fibre-diagram of structure B, we may state the following conclusions. The structure is probably helical. The phosphate groups lie on the outside of the structural unit, on a helix of diameter about 20 A. There are 10 phosphate groups per chain in one turn of the helix. The structure does not contain more than one equivalent co-axial chain, but the possibility of non-equivalent co-axial chains is not eliminated.

The total absence of an inner maximum on the fourth layer-line suggests that if there are 2 non-equivalent co-axial chains these are separated by $\frac{3}{8}$ of the fibre-axis period, that is by $\frac{3}{8} \times 41$ A in the fibre-axis direction.

≈ 13 A.